Raw Sequence Listing Error Summary

ERROR DETECTED SUGGESTED CORRECTION SERIAL NUMBER: 09/201,916

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 _____ Wrapped Nucleics
The number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

2 _____ Wrapped Aminos
The amino acid number/text at the end of each line "wrapped " down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

3 _____ Incorrect Line Length
The rules require that a line not exceed 72 characters in length. This includes spaces.

4 _____ Misaligned Amino Acid Numbering
The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.

5 _____ Non-ASCII
This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6 _____ Variable Length
Sequence(s) ____ contain n's or Xaa's which represented more than one residue.
As per the rules, each n or Xaa can only represent a single residue.
Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.

7 _____ PatentIn ver. 2.0 "bug"
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence.

8 _____ Skipped Sequences (OLD RULES)
Sequence(s) ____ missing. If intentional, please use the following format for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X:
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9 _____ Skipped Sequences (NEW RULES)
Sequence(s) ____ missing. If intentional, please use the following format for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

10 _____ Use of n's or Xaa's (NEW RULES)
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

11 _____ Use of <213>Organism (NEW RULES)
Sequence(s) _____ are missing this mandatory field or its response.

12 _____ Use of <220>Feature (NEW RULES)
Sequence(s) ____ are missing the <220>Feature and associated headings.
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"
Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)

13 _____ PatentIn ver. 2.0 "bug"
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other means to copy file to floppy disk.

AKS-Biotechnology Systems Branch- 5/15/99

R. Zeman

RAW SEQUENCE LISTING
PATENT APPLICATION **US/09/201,916**

DATE: 12/30/1999
TIME: 15:16:58

Input Set: **I201916.RAW**

This Raw Listing contains the General Information Section and up to first 5 pages.

Does Not Comply
Corrected Diskette Needed
pp 3-4

<110> APPLICANT: Hope, Ralph Graham
Mclauchlan, John
<120> TITLE OF INVENTION: VIRAL THERAPEUTICS
<130> FILE REFERENCE: DYOU17.001AUS
<140> CURRENT APPLICATION NUMBER: US/09/201,916
<141> CURRENT FILING DATE: 1998-12-01
<160> NUMBER OF SEQ ID NOS: 13
<170> SOFTWARE: FastSEQ for Windows Version 3.0
<210> SEQ ID NO 1
<211> LENGTH: 630
<212> TYPE: DNA
<213> ORGANISM: Hepatitis C Virus
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (43)...(630)
<400> SEQUENCE: 1

```
ggtgcttgcg agtgccccgg gaggtctcgt agaccgtgca cc atg agc acg aat        54
                                                 Met Ser Thr Asn
                                                   1

cct aaa cct caa aga aaa acc aaa cgt aac acc aac cgt cgc cca cag      102
Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln
 5                  10                  15                  20

gac gtt aag ttc ccg ggt ggc ggt cag atc gtt ggt gga gtt tac ttg      150
Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr Leu
                25                  30                  35

ttg ccg cgc agg ggc cct aga ttg ggt gtg cgc gcg acg agg aag act      198
Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala Thr Arg Lys Thr
            40                  45                  50

tcc gag cgg tcg caa cct cga ggt aga cgt cag cct atc ccc aag gca      246
Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro Ile Pro Lys Ala
        55                  60                  65

cgt cgg ccc aag ggc agg aac tgg gct cag ccc ggg tat cct tgg ccc      294
Arg Arg Pro Lys Gly Arg Asn Trp Ala Gln Pro Gly Tyr Pro Trp Pro
    70                  75                  80

ctc tat ggc aat gag ggt tgc ggg tgg gcg gga tgg ctc ctg tcc ccc      342
Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp Leu Leu Ser Pro
 85                  90                  95                 100

agt ggc tct cgg cct agt tgg ggc ccc aac gac ccc cga cgt agg tcg      390
Ser Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro Arg Arg Arg Ser
                105                 110                 115

cgc aat ttg ggt aag gtc atc gat acc ctt acg tgc ggc ttc gtc gat      438
Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Val Asp
            120                 125                 130

ctc atg ggg tac ata ccg ctc gtc ggc gcc cct ctt aga ggc gct gcc      486
```

```
45          Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Arg Gly Ala Ala
46                  135                 140                 145
47          agg gcc ctg gcg cat ggc gtc cgg gtt ctg gaa gac ggt gtg aac tat      534
48          Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr
49              150                 155                 160
50          gca aca ggt aac ctt cct ggt tgc tct ttc tct atc ttc ctt ctg gcc      582
51          Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala
52          165                 170                 175                 180
53          ctg ctc tct tgc ctg act gtg ccc gct tca gcc tac caa gtg cgc aac      630
54          Leu Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Gln Val Arg Asn
55                      185                 190                 195
56      <210> SEQ ID NO 2
57      <211> LENGTH: 60
58      <212> TYPE: DNA
59      <213> ORGANISM: Hepatitis C Virus
60      <220> FEATURE:
61      <221> NAME/KEY: CDS
62      <222> LOCATION: (1)...(60)
63      <223> OTHER INFORMATION: Corresponds to aa 125 to 144 of SEQ ID. No. 1
64      <400> SEQUENCE: 2
65          acc ctt acg tgc ggc ttc gtc gat ctc atg ggg tac ata ccg ctc gtc       48
66          Thr Leu Thr Cys Gly Phe Val Asp Leu Met Gly Tyr Ile Pro Leu Val
67           1               5                  10                  15
68          ggc gcc cct ctt                                                      60
69          Gly Ala Pro Leu
70                      20
71      <210> SEQ ID NO 3
72      <211> LENGTH: 18
73      <212> TYPE: DNA
74      <213> ORGANISM: Hepatitis C Virus
75      <220> FEATURE:
76      <221> NAME/KEY: CDS
77      <222> LOCATION: (1)...(18)
78      <223> OTHER INFORMATION: Corresponds to aa 161-166 of SEQ ID. No. 1
79      <400> SEQUENCE: 3
80          ggt gtg aac tat gca aca                                               18
81          Gly Val Asn Tyr Ala Thr
82           1               5
83      <210> SEQ ID NO 4
84      <211> LENGTH: 1900
85      <212> TYPE: DNA
86      <213> ORGANISM: Human
87      <400> SEQUENCE: 4
88          cgtcttcggg acgcgcccgc tcttcgcctt tcgctgcagt ccgtcgattt ctttctccag      60
89          gaagaaaaat ggcatccgtt gcagttgatc cacaaccgag tgtggtgact cgggtggtca     120
90          acctgccctt ggtgagctcc acgtatgacc tcatgtcctc agcctatctc agtacaaagg     180
91          accagtatcc ctacctgaag tctgtgtgtg agatgscaga gaacggtgtg aagaccatca     240
92          cctccgtggc catgaccagt gctctgccca tcatccagaa gctagagccg caaattgcag     300
93          ttgccgatac ctatgcctgt aaggggctag acaggattga ggagagactg cctattctga     360
94          atcagccatc aactcagatt gttgccaatg ccaaaggcgc tgtgactggg gcaaaagatg     420
```

W--> 95 ctgtgacgac tactgtgact ggggccaagg attctgtngc cagcacgatc acaggggtga 480
96 tggacaagac caaaggggca gtgactggca gtgtggagaa gaccaagtct gtggtcagtg 540
97 gcagcattaa cacagtcttg gggagtcgga tgatgcagct cgtgagcagt ggcgtagaaa 600
98 atgcactcac caaatcagag ctgttggtag aacagtacct ccctctcact gaggaagaac 660
99 tagaaaaaga agcaaaaaaa gttgaaggat ttgatctggt tcagaagcca agttattatg 720
100 ttagactggg atccctgtct accaagcttc actcccgtgc ctaccagcag gctctcagca 780
101 gggttaaaga agctaagcaa aaaagccaac agaccatttc tcagctccat tctactgttc 840
102 acctgattga atttgccagg aagaatgtgt atagtgccaa tcagaaaatt caggatgctc 900
103 aggataagct ctacctctca tgggtagagt ggaaaaggag cattggatat gatgatactg 960
104 atgagtccca ctgtgctgag cacattgagt cacgtactct tgcaattgcc cgcaacctga 1020
105 ctcagcagct ccagaccacg tgccacaccc tcctgtccaa catccaaggt gtaccacaga 1080
106 acatccaaga tcaagccaag cacatggggg tgatggcagg cgacatctac tcagtgttcc 1140
107 gcaatgctgc ctcctttaaa gaagtgtctg acagcctcct cacttctagc aaggggcagc 1200
108 tgcagaaaat gaaggaatct ttagatgacg tgatggatta tcttgttaac aacacgcccc 1260
109 tcaactggct ggtaggtccc ttttatcctc agctgactga gtctcagaat gctcaggacc 1320
110 aaggtgcaga gatggacaag agcagccagg agacccagcg atctgagcat aaaactcatt 1380
111 aaacctgccc ctatcactag tgcatgctgt ggccagacag atgacacctt ttgttatgtt 1440
112 gaaattaact tgctaggcaa ccctaaattg ggaagcaagt agctagtata aaggccctca 1500
113 attgtagttg tttccagctg aattaagagc tttaaagttt ctggcattag cagatgattt 1560
114 ctgttcacct ggtaagaaaa gaatgatagg cttgtcagag cctatagcca gaactcagaa 1620
115 aaaattcaaa tgcacttatg ttctcattct atggccattg tgttgcctct gttactgttt 1680
116 gtattgaata aaaacatctt catgtgggct ggggtagaaa ctggtgtctg ctctggtgtg 1740
117 atctgaaaag gcgtcttcac tgctttatct catgatgctt gcttgtaaaa cttgatttta 1800
118 gtttttcatt tctcaaatag gaatactacc tttgaattca ataaaattca ctgcaggata 1860
W--> 119 gaccagttna gnagcaaaca nncangtaca cnnaaganac 1900

120 <210> SEQ ID NO 5
121 <211> LENGTH: 437
122 <212> TYPE: PRT
123 <213> ORGANISM: Human
124 <400> SEQUENCE: 5
125 Met Ala Ser Val Ala Val Asp Pro Gln Pro Ser Val Val Thr Arg Val
126 1 5 10 15
127 Val Asn Leu Pro Leu Val Ser Ser Thr Tyr Asp Leu Met Ser Ser Ala
128 20 25 30
129 Tyr Leu Ser Thr Lys Asp Gln Tyr Pro Tyr Leu Lys Ser Val Cys Glu
130 35 40 45
W--> 131 Met Xaa Glu Asn Gly Val Lys Thr Ile Thr Ser Val Ala Met Thr Ser
132 50 55 60
133 Ala Leu Pro Ile Ile Gln Lys Leu Glu Pro Gln Ile Ala Val Ala Asp
134 65 70 75 80
135 Thr Tyr Ala Cys Lys Gly Leu Asp Arg Ile Glu Glu Arg Leu Pro Ile
136 85 90 95
137 Leu Asn Gln Pro Ser Thr Gln Ile Val Ala Asn Ala Lys Gly Ala Val
138 100 105 110
139 Thr Gly Ala Lys Asp Ala Val Thr Thr Thr Val Thr Gly Ala Lys Asp
140 115 120 125
141 Ser Val Ala Ser Thr Ile Thr Gly Val Met Asp Lys Thr Lys Gly Ala
142 130 135 140
143 Val Thr Gly Ser Val Glu Lys Thr Lys Ser Val Val Ser Gly Ser Ile
144 145 150 155 160

all item 10 on Error Summary Sheet

item 10

Asn Thr Val Leu Gly Ser Arg Met Met Gln Leu Val Ser Ser Gly Val
165 170 175
Glu Asn Ala Leu Thr Lys Ser Glu Leu Leu Val Glu Gln Tyr Leu Pro
180 185 190
Leu Thr Glu Glu Glu Leu Glu Lys Glu Ala Lys Lys Val Glu Gly Phe
195 200 205
Asp Leu Val Gln Lys Pro Ser Tyr Tyr Val Arg Leu Gly Ser Leu Ser
210 215 220
Thr Lys Leu His Ser Arg Ala Tyr Gln Gln Ala Leu Ser Arg Val Lys
225 230 235 240
Glu Ala Lys Gln Lys Ser Gln Gln Thr Ile Ser Gln Leu His Ser Thr
245 250 255
Val His Leu Ile Glu Phe Ala Arg Lys Asn Val Tyr Ser Ala Asn Gln
260 265 270
Lys Ile Gln Asp Ala Gln Asp Lys Leu Tyr Leu Ser Trp Val Glu Trp
275 280 285
Lys Arg Ser Ile Gly Tyr Asp Asp Thr Asp Glu Ser His Cys Ala Glu
290 295 300
His Ile Glu Ser Arg Thr Leu Ala Ile Ala Arg Asn Leu Thr Gln Gln
305 310 315 320
Leu Gln Thr Thr Cys His Thr Leu Leu Ser Asn Ile Gln Gly Val Pro
325 330 335
Gln Asn Ile Gln Asp Gln Ala Lys His Met Gly Val Met Ala Gly Asp
340 345 350
Ile Tyr Ser Val Phe Arg Asn Ala Ala Ser Phe Lys Glu Val Ser Asp
355 360 365
Ser Leu Leu Thr Ser Ser Lys Gly Gln Leu Gln Lys Met Lys Glu Ser
370 375 380
Leu Asp Asp Val Met Asp Tyr Leu Val Asn Asn Thr Pro Leu Asn Trp
385 390 395 400
Leu Val Gly Pro Phe Tyr Pro Gln Leu Thr Glu Ser Gln Asn Ala Gln
405 410 415
Asp Gln Gly Ala Glu Met Asp Lys Ser Ser Gln Glu Thr Gln Arg Ser
420 425 430
Glu His Lys Thr His
435

<210> SEQ ID NO 6
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: A branched peptide containing residues 5-27 of the HCV core protein with ~~degeneracy at positions 1 and 12 in which~~ position 1 can be Ala or Pro and position 12 can be Ile or Asn.

use Xaa and explain in (2207-(2237 section

<400> SEQUENCE: 6
Ala Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Ile Arg Arg Pro Gln
1 5 10 15
Asp Val Lys Phe Pro Gly Gly Lys Lys Lys Lys Lys Lys Ala
20 25 30

<210> SEQ ID NO 7
<211> LENGTH: 11
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotides used to construct HCV core protein deletion plasmids.
<400> SEQUENCE: 7
gctgagatct a 11

<210> SEQ ID NO 8
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotides used to construct HCV core protein deletion plasmids.
<400> SEQUENCE: 8
gtaaccttcc tggttgctct tgagatcta 29

<210> SEQ ID NO 9
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotides used to construct HCV core protein deletion plasmids.
<400> SEQUENCE: 9
gtaacctttg agatcta 17

<210> SEQ ID NO 10
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotides used to construct HCV core protein deletion plasmids.
<400> SEQUENCE: 10
ctggcgcatt gagatcta 18

<210> SEQ ID NO 11
<211> LENGTH: 28
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotides used to construct HCV core protein deletion plasmids.
<400> SEQUENCE: 11
ctggcccatg gtgttaacta tgcaacag 28

<210> SEQ ID NO 12
<211> LENGTH: 31
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:

VERIFICATION SUMMARY
PATENT APPLICATION US/09/201,916

DATE: 12/30/1999
TIME: 15:16:58

Input Set: I201916.RAW

Line	?	Error/Warning	Original Text
95	W	"N" or "Xaa" used: Feature required	ctgtgacgac tactgtgact ggggccaagg attctgtn
119	W	"N" or "Xaa" used: Feature required	gaccagttna gnagcaaaca nncangtaca cnnaagan
131	W	"N" or "Xaa" used: Feature required	Met Xaa Glu Asn Gly Val Lys Thr Ile Thr S